AF
IFW

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Applicant : NEVENKA DIMITROVA ET AL. Examiner: TING ZHOU

Serial No.: 09/866,394 Group Art Unit: 2173

Filed : May 25, 2001

For : COMPACT VISUAL SUMMARIES USING SUPERHISTOGRAMS AND FRAME SIGNATURES.

Board of Patent Appeals and Interferences
United States Patent and Trademark Office
PO Box 1450
Alexandria, VA 22313-1450

Sir:

Enclosed is an Appeal Brief in the above-identified patent application.

Please charge the fee of $500.00 as indicated on Form No. 2038.

Respectfully submitted,

Michael A. Scaturro
Reg. No. 51,356
Attorney for Applicant
(516) 414-2007

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited this date with the United States Postal Service as first-class mail in an envelope addressed to:

COMMISSIONER FOR PATENTS
PO Box 1450
Alexandria, VA 22313-1450

On 7-25-2005
(Date of Mailing)

By ______
(Signature)

PATENT
Atty. Docket: [MS-164] PHUS-0100265

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant	: NEVENKA DIMITROVA ET AL.	Examiner: TING ZHOU
Serial No.	: 09/866,394	Group Art Unit: 2173
Filed	: May 25, 2001	
For	: COMPACT VISUAL SUMMARIES USING SUPERHISTOGRAMS AND FRAME SIGNATURES.	

Board of Patent Appeals and Interferences
United States Patent and Trademark Office
PO Box 1450
Alexandria, VA 22313-1450

APPEAL BRIEF

Sir:

Appellant herewith respectfully presents its Brief on Appeal as follows:

REAL PARTY IN INTEREST

The real party in interest is Koninklijke Philips Electronics N.V., a corporation of The Netherlands having an office and a place of business at Groenewoudseweg 1, Eindhoven, Netherlands 5621 BA. Koninklijke Philips Electronics N.V. is the parent company of the assignee of record U.S. Philips Corporation, a Delaware corporation having an office and a place of business at 345 Scarborough Road, Briarcliff Manor, New York, 10510-8001.

RELATED APPEALS AND INTERFERENCES

To the best of Appellants' knowledge and belief, there are no related appeals or interferences.

STATUS OF CLAIMS

Claims 1-5, 7-15, 17- 24, 26-33 and 35-38 are pending in this application. Claims 1-5, 7-15, 17- 24, 26-33 and 35-38 are rejected in the Final Office Action that mailed January 5, 2005. This rejection was upheld in an Advisory Action that mailed April 19, 2005. Claims 1-5, 7-15, 17- 24, 26-33 and 35-38 are the subject of this appeal. A copy of claims 1-5, 7-15, 17- 24, 26-33 and 35-38 are presented in Appendix A.

STATUS OF AMENDMENTS

An Amendment after Final Action was filed March 25, 2005 in response to the Final Office Action. The Advisory Action upheld the rejection in response to that amendment. This Appeal Brief is in response to the Final Office Action that rejected Claims 1-5, 7-15, 17- 24, 26-33 and 35-38 and the Advisory Action that upheld that rejection.

SUMMARY OF CLAIMED SUBJECT MATTER

A first aspect of the present invention, for example as claimed in independent Claim 1 relates to an apparatus for use in a system capable of creating visual summaries of video material, as described in the specification such as page 4, lines 6-15 and at page 7, line 14 through page 15, line 17. A second aspect of the invention provides a system capable of creating visual summaries of video material, as described in the specification such as page 7, line 14 through page 15, line 17. A third aspect of the invention provides a method of locally enhancing display information, for example as claimed in independent claim 21 and as described in the specification such as page 15, line 18 through page 16, line 8. A fourth aspect of the present invention provides computer executable instructions capable of creating visual summaries of video material, for example as claimed in independent Claim 31 and as described in the specification such as page 9, lines 15 – 19.

The apparatus and system, as shown in FIG. 1 of the specification, and as described in the specification such as page 4, lines 6-15 and at page 7, line 14 through page 15, line 17, includes a visual summary controller 130 comprised of a keyframe filter module 140, a color information module 150, a histogram and keyframe selection module 160, a visual summary module 170 and a visual summary retrieval module 180.

The method of the invention, as claimed in claim 21, includes the step of the controller 130 receiving keyframes from the video processor 110 (step 405). The method further includes the step of the controller 130 extracting frame signatures from the keyframes and filtering the keyframes (step 410). The method then describes the step of the controller 130 deriving color information from the filtered keyframes (step 415). Next, the method describes the step of the controller 130 deriving superhistograms from the color information (step 420). Next, the controller 130 selects a representative keyframe or a representative set of multiple keyframes for each family histogram (step 425). At a next step, the controller 130 creates a compact visual summary from the selected keyframe images (step 430). Next, the controller 130 stores the compact visual summary in a visual summary storage location 270 within memory unit 120 (step 435). When requested by a user, visual summary retrieval module 180 retrieves a visual summary from memory unit 120 and causes it to be displayed (step 440).

GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Whether Claims 1-5, 7-15, 17- 24, 26-33 and 35-38 are unpatentable over the article entitled "Color SuperHistograms for Video Representation", written by Dimitrova et al. in view of U.S. Patent No. 5,805,733 issued to Wang et al. on September 8, 1998 ("Wang et al."). The Appellants respectfully request the Board to address the patentability of independent claims 1, 11, 21, and 30, based on the requirements of Claim 1. This position is provided for the specific purpose and stated purpose of simplifying the current issue on appeal. However, the Appellants herein specifically reserve the right to argue and address the patentability of each of the further claims at a later date should the separately patentable subject matter of those claims at a later date should the separately patentable subject matter of those claims later become an issue. Accordingly, this limitation of the subject matter presented for appeal herein, specifically limited to discussions of the patentability of claims 1, 11, 21, and 30 is not intended as a waiver of Appellants' right to argue the patentability of the further claims and claim elements at that later time.

ARGUMENT

Claims 1, 11, 21, and 30 are said to be unpatentable over the article entitled "Color SuperHistograms for Video Representation", written by Dimitrova et al. (hereafter "Dimitrova") in view of U.S. Patent No. 5,805,733 issued to Wang et al. on September 8, 1998 (hereafter "Wang").

The Examiner states in the Final Office Action, mailed on January 25, 2005, that the Dimitrova article teaches an apparatus, system, method and computer executable instructions comprising a visual summary controller capable of creating a visual summary of video material, wherein the visual summary controller is capable of extracting frame signatures (histograms) from keyfreames of video material and capable of using the frame signatures to create superhistograms from the keyframes. The Appelants agree with the Examiner's assertion with regard to Dimitrova.

The Examiner, by admission, states that the Dimitrova article fails to explicitly teach an element of Claim 1 that recites:

> **...selecting representative keyframe images for each superhistogram to create a compact visual summary of the video material, wherein the representative images include at least one of**
>
> **(a) the first image in each family histogram,**
>
> **(b) the most meaningful image in each superhistogram,**
>
> **(c) a randomly chosen image, and**
>
> **(d) an image that is closest to the cluster center.**

The Examiner states that Wang et al. teaches the analysis of scenes and frames in video materials (Wang et al. column 1, lines 53-56 and Figure 2) similar to that of Dimitrova et al. In addition, Wang et al. further teaches selecting representative keyframe images from each group of related scenes to create a compact visual summary of the video material (summarizing a video sequence by taking one representative frame from each set of related scenes with similar average color histograms, to represent the set to enable the user to view a large sampling of video sequence images) (Wang et al: column 1, lines 51-67 and column 2, lines 1-24; this is further shown in Figure 3), wherein the representative images include at least one of (a) the first image in each family histogram, **(b) the most meaningful image in each superhistogram,** (c) a randomly chosen image, and **(d) an image that is closest to the cluster center (Wang et al: column 3, lines 37-66).**

The Examiner asserts that it would have been obvious to modify the visual summary controller capable of extracting frame signatures from keyframes to create superhistograms, as taught in Dimirova, to include the further step of selecting representative keyframes from those superhistograms and using the representative keyframe images to create a compact visual summary, taught by Wang. In the "Response to Arguments", the Examiner states that it can be seen that Wang et al. teaches that the representative images includes **at least one of** the first image in each family histogram, **the most meaningful image in each superhistogram**, a randomly chosen image and **an image that is closest to the cluster center.** In the Office Action, the Examiner specifically highlights the terms **the most meaningful image in each superhistogram**, and **an image that is closest to the cluster center** as the two terms from among the four terms recited which are allegedly taught by Wang.

Appelants' position

The Examiner is incorrect in her assertion that Wang teaches: ...selecting representative keyframe images for each **superhistogram** to create a compact visual summary of the video material, wherein the representative images include at least one of (b) **the most meaningful image in each superhistogram**, (d) **an image that is closest to the cluster center**.

The Appelants' previously argued in the Final Office Action that Wang's method for selecting representative frames is based on a temporal ordering of frames. In contrast, the method of the invention selects representative frames based on a non-temporal ordering. Frame selection based on a non-temporal ordering is supported by the specification and though the use of certain terms used throughout the claims. The claim terminology that is indicative of a non-temporal ordering include terms defined in the specification by the Appelants as lexicographer. With specific reference to Claim 1, the terms that are indicative of a non-temporal ordering include: 'superhistogram', 'family histogram', and 'cluster center'.

While Applicants readily acknowledge that the Examiner must give the pending claims their broadest interpretation, consistent with the specification, the Appelants submit that the law is clear that when the Appelants choose to define terms as lexicographers in the specification it is incumbent upon the Examiner to analyze the claim language in light of their ascribed definitions in the specification in order to achieve a complete exploration of the applicant's invention and its relation to the prior art. Support is found in the MPEP at 2173.05(a), where it states

MPEP 2173.05(a)

The meaning of every term used in a claim should be apparent from the prior art or from the specification and drawings at the time the application is filed. Applicants need not confine themselves to the terminology used in the prior art, but are required to make clear and precise the terms that are used to define the invention whereby the metes and bounds of the claimed invention can be ascertained. During patent examination, the pending claims must be given the broadest reasonable interpretation consistent with the specification. *In re Morris*, 127 F.3d 1048, 1054, 44 USPQ2d 1023, 1027 (Fed. Cir. 1997); *In re Prater*, 415 F.2d 1393, 162 USPQ 541 (CCPA 1969). See also MPEP § 2111 - § 2111.01. When the specification states the meaning that a term in the claim is intended to have, the claim is examined using that meaning, in order to achieve a complete exploration of the applicant's invention and its relation to the prior art. In re Zletz, 893 F.2d 319, 13 USPQ2d 1320 (Fed. Cir. 1989).

The Wang Patent

The Wang patent discloses a method and system for detecting scenes and summarizing a video sequence or any other temporally ordered sequence of images into a number of distinct scenes.

The principle employed by Wang is that similar scenes will have substantially similar average color distributions. Each group of related scenes is then represented by a frame selected from the set of scenes by displaying the representative frame to the user – see Wang in the Summary at Col. 1, lines 50-67.

Wang further discloses at Col. 2, lines 12-15 that, "a representative frame is then taken for each group of related scenes. The representative frames can be a frame that is halfway between first and last scenes in the group of redundant scenes."

In the Final Office Action, the Examiner states at page 7 in the "**Response to Arguments**", that Wang teaches selecting a representative frame image for each set of summarized related scenes, as recited in Col. 3, lines 37-57:

Referring now to FIG. 2, there is shown a flowchart for the method of summarizing video sequences. The user inputs 201 a video sequence into the system 100, or retrieves a video sequence from the mass storage device 107. The scene change detector 121 then processes the video sequence to detect 203 individual scenes. The related scene detector identifies 205 disparate related scenes, consolidating the video sequence into a smaller number of scenes. The related scenes are then time ordered 207, and displayed 209 by the user interface controller 125, showing a representative frame from each set of related scenes. An index of the scenes in the video sequence is also displayed 211, here the movie bar is created by the movie bar generator 127. The user may then select 213 any number of scenes for viewing in their entirety, the system 100 retrieving 215 each selected scene, and displaying it to the user.

FIG. 3 shows an illustration of one embodiment of a user interface for displaying summarized scenes, as produced by the user interface controller 125. A window 301 includes a collage 303 made up of a representative frame 305 for each set of summarized scenes.

While Wang may arguably teach the selection of a representative frame image for each set of summarized related scenes, as recited in Col. 3, lines 37-57, it is respectfully submitted, however, that Wang does not teach that the representative image includes **at least one of** the first image in each family histogram, **the most meaningful image in each superhistogram**, a randomly chosen image and **an image that is closest to the cluster center**, as asserted by the Examiner in section 4, "**Response to Arguments**", in the Final Office Action. The Examiner provides explicit support at least for the assertion that Wang teaches that the representative image includes at least one of **the most meaningful image in each superhistogram** and **an image that is closest to the cluster center** as follows.

The Examiner states that Wang et al. teaches that the representative frame can be a frame that is halfway between the first and last scenes, as recited in column 2, lines 12-15 and column 3, lines 57-59, In other words, the representative frame can include a frame that is in **the middle, or center, of the cluster of scenes**.

The Examiner further states that Wang et al. also teach that the representative frame can be a frame taken from the longest scene, since the longest scene is most indicative of the content of the related scenes, as recited in column 3, lines 59-62, in other words, the representative frame can include a frame that is the most indicative of the contents of the related scenes, i.e., **the most meaningful frame in the group**.

An example is provided below to more clearly illustrate why Appelants believe that Wang does not teach that the representative image includes **at least one of the most meaningful image in each superhistogram**, and **an image that is closest to the cluster center**, as asserted by the Examiner.

Example

Assume a video sequence is input by a user, the video sequence being comprised of a number of frames, e.g., frames 1, 2, 3, ... , 309, 310, 311.

Wang teaches a method for summarizing the video sequence in the flowchart of Fig. 2. According to Wang, a scene change detector processes the video sequence to detect individual scene changes. In the example, it is assumed that the scene change detector outputs six scenes from the exemplary input sequence of 311 frames.

Scene 1 – made up of frames 1 – 20
Scene 2 - made up of frames 21 – 77
Scene 3 – made up of frames 78 – 160
Scene 4 – made up of frames 161 – 203
Scene 5 - made up of frames 204 – 255
Scene 6 - made up of frames 256 – 311

The video sequence (1-311) is thus shown to be consolidated into a number of scenes. According to Wang, related scenes from among the six detected scenes are then time ordered. In the example, it is assumed that scenes 1, 3 and 6 constitute a first set of related scenes, scenes 2 and 5 constitute a second set of related scenes and scene 4 constitutes is unrelated to all other scenes.

<u>First Set of Related Scenes in Time order</u>

<u>Scene 1</u> – made up of frames 1 – 20

<u>Scene 3</u> – made up of frames 78 – 160

Scene 6 - made up of frames 256 – 311

<u>Second Set of Related Scenes in Time order</u>

<u>Scene 2</u> – made up of frames 21 – 77

<u>Scene 5</u> – made up of frames 204 – 255

Third Set of Related Scenes in Time order

Scene 4 – made up of frames 161 – 203

Wang teaches that the respective time ordered sets of related scenes are displayed to a user showing a frame from each set of related scenes. In the instant example, according to Wang, 1 frame is shown from the first set, 1 frame is shown from the second set and 1 frame is shown from the third set.

In the Final Office Action, the Examiner asserts that Wang et al. teaches that a representative keyframe can be **an image that is closest to the cluster center** by stating that the representative frame can be a frame that is halfway between the first and last scenes, as recited in column 2, lines 12-15 and column 3, lines 57-59.

Wang teaches at Col. 6, lines 9-25, that a frame that is halfway between the first and last scenes may be computed as a frame **Fmid,** selected as the mid-point of all the related scenes or may otherwise be selected as the middle frame of the longest scene from among the related scenes.

> This process is repeated 507 for each window of n scenes, beginning with the first scene that was selected 503. In the preferred embodiment, the window is "advanced" by one scene each time, though in alternate embodiments, a larger step may be taken between windows. After each window of scenes has been analyzed, and all related scenes identified, then 517 in each set of related scenes, the total scene time is determined 519, and the frame **Fmid** that is the midpoint of all the scenes is chosen 521 as a representative frame for all the related scenes. Alternatively, the middle frame of the longest scene in each set of related scenes can be used as **Fmid**. Referring to FIG. 2 again, the related scenes are time ordered 207, and the **Fmid** frames for each set of scenes are displayed 209 to the user by the user interface controller 125. The user may view a scene by selecting 213 one of the representative frames with the pointing device 105, the system 100 retrieving 215 and displaying the scenes associated with the representative frame.

The calculation of **Fmid** according to Wang is now performed in the context of the instant example. Using the first set of related scenes for purposes of illustration,

The total duration of the three scenes is computed as 157:

157 = 20 (from scene 1) + 82 (from scene 2) + 55 (from scene 3)

where, the midpoint may then be computed as:

78 = 150 / 2 = the midpoint

As a consequence of Wang being subject to more than one interpretation in this regard, an alternative calculation of the midpoint can be made by considering the beginning of the related scenes to the end of the related scenes:

155 = (311 – 1) / 2

It is therefore shown, by way of example, that a frame selected as **an image that is closest to the cluster center** is computed in accordance with the method of Wang as one of frame 78 or frame 157, depending upon the reader's interpretation.

Method of the Invention

The Appellants assert that that the method according to the invention for calculating a representative keyframe image that is **the most meaningful frame in the group** or **an image that is closest to the cluster center,** will produce results that are in sharp contrast to those produced according to the method of Wang. The differences are made apparent in light of the instant example.

As stated above, a key distinction between Wang and the invention is that Wang's method for selecting representative frames is based on a temporal ordering of frames. In sharp contrast, the method of the invention for selecting representative keyframe images is based on a non-temporal ordering of frames. This significant difference is evidenced in the claim language and throughout the specification whereby the meaning of certain terms used in the claims, by virtue of their ascribed definitions, yielding results which are in sharp contrast to the results achieved in the prior art. Accordingly, as argued in the Final Office Action, it is incumbent upon the Examiner to analyze the claim language in light of their ascribed definitions in the specification in order to achieve a complete exploration of the applicant's invention and its relation to the prior art.

In the Final Office Action, the Appelants attempted to show, by way of example, that the use of particular terms in the claims (i.e., superhistogram, family histogram, cluster center), yield results for selecting representative keyframes corresponding to (a) **the most meaningful frame in the group,** and (2) **an image that is closest to the cluster center,** which are appreciably different from the results obtained according to the method of Wang.

The following tables are provided in further support of the instant example to more clearly illustrate these differences.

Table I.

	Scene label and histogram of the representative keyframe	Representative Keyframe	Family
Scene 1 – made up of frames 1-20	(outdoor1 scene) [4 5 4 5 2]	1	1
Scene 2 – made up of frames 21-77	(indoor1 scene4) - [5 2 2 0 11] - [5 2.5 2 .5 10]	21	2
Scene 3 – made up of frames 78-160	(outdoor1 scene) - [5 4 5 3 3]	78	1
Scene 4 – made up of frames 161-203	(outdoor2 scene) - [8 1 6 2 3]	161	3
Scene 5 – made up of frames 204-255	(indoor1 scene) - [5 3 2 1 9]	204	2
Scene 6 – made up of frames 256-311	(outdoor1 scene) - [4 5 3 6 2]	256	1

Table II.

	Scenes	Representative keyframe picked by cut detection	Cumulative histogram	Histogram Distances to the cumulative histogram	Frame selected Dimitrova				Frame selected Wang	
					1 *	2 *	3 *	4*	1	2
Family	1, 3, 6	1, 78, 256	[4.3 ,4.6 ,4 ,4.6,	(1.4, 4.6, 3.8)	1	256	78	1	136	119

1			2.3]						or 155	
Family 2	2, 5	21, 204	[5, 2.5, 2, .5, 10]	(2, 2)	21	204	21	21	74 or 91	49
Family 3	4	161	[8, 1, 6, 2, 3]	(0)	161	161	161	161	182	283

Where 1 = 1st image in each family histogram
2 = the most meaningful image in each family histogram (the example assumes frame 256 contains a face)
3 = a randomly chosen image
4 = an image that is closest to the cluster center

For ease of explanation, reference will focus exclusively on Family 1 (i.e. related scenes 1. 3 and 6). According to principles of the invention, a cumulative histogram is computed from the histograms of the representative keyframes of scenes 1, 3 and 6. The cumulative histogram is a construct that is a product of the family histograms and superhistograms, discussed throughout the specification and as used in the claims.

The cumulative histogram for Family 1 is shown in Table II as [4.3, 4.6, 4, 4.6, 2.3]. The cumulative histogram values are derived by averaging the histograms for the representative keyframes in each of scenes 1, 3 and 6. For example, the first value, 4.3, is derived by averaging the first histogram value in each of scenes 1, 3 and 6.

$$4.3 = (4 + 5 + 4) / 3$$

The other values in the cumulative histogram are derived in like manner. A critical distinction between the method of Wang and the method of the invention is that, in accordance with the method of the invention, the **cluster center is synonymous with the cumulative histogram**. Whereby the cluster center is derived as a non-temporal ordering.

In contrast to the method of the invention, Wang does not teach or disclose the computation of a **cluster center** in accordance with the principles of the invention. Rather, Wang computes a cluster center value using Fmid, as described above.

In accordance with the method of the invention, **the image that is closest to the cluster center** can be calculated for Family 1 in the following way. For Family 1, the histogram distances to the cluster center for each of the respective representative keyframes images from scenes 1, 3 and 6 are computed as (1.4, 4.6, 3.8), respectively. The <u>minimum</u> distance to the cluster center is therefore shown to be 1.4, associated with representative keyframe 1 of scene 1. In general, the distances from a keyframe's histogram to the cluster center is computed by taking the absolute difference in the histogram valuesand summing the result. As an example, the distance computation for keyframe 1 (i.e., the minimum distance) is as follows.

$$1.4 = ABS(4.3 - 4) + ABS(5-4.6) + ABS(4-4) + ABS(5-4.6) + ABS(2-2.3)$$

$$= .3 + .4 + 0 + .4 + .3$$

It is therefore shown that keyframe 1 having a minimum distance of 1.4, qualifies the keyframe as being **closest to the cluster center**. This is in sharp contrast with the selected values keyframe values of 78 and 155, selected as being **closest to the cluster center** in accordance with the method of Wang.

Based on the above, it is therefore respectfully submitted that Wang does not teach or disclose, "wherein the representative images include at least one of .... **an image that is closest to the cluster center".**

In the final office action, the Examiner also asserts that Wang et al. teaches **an image that is the most meaningful frame in the group** by asserting that the representative frame can be a frame that is taken from the longest scene, since the longest scene is most indicative of the content of the related scenes, as recited in column 3, lines 59-62.

The specification recites at page 14, line 13, " The term "most meaningful image" may refer to a frame with a person's face, an important text". In the final office action, in the "Response to Arguments", the Examiner asserts that Wang also teaches that the representative frame can be a frame taken from the longest scene, since the longest scene is most indicative of the contents of the related scenes. The Examiner asserts that the representative frame can include a frame that is the most indicative of the contents of the related scenes, i.e., the most meaningful in the group. The Appelants respectfully disagree. The specification recites with particularity (by example) what constitutes the 'most meaningful frame", i.e., a person's face, an important text. It is respectfully submitted that simply selecting one frame from among the frames of the longest scene, as taught in Wang, does not teach or disclose the "most meaningful frame in the group", as recited in Claim 1.

CONCLUSION

Claims 1-5, 7-15, 17-24, 26-33 and 35-38 are patentable over Wang.

Thus the Examiner's rejection of Claims 1-5, 7-15, 17-24, 26-33 and 35-38 should be reversed.

Respectfully submitted,

Michael A. Scaturro
Reg. No. 51,356
Attorney for Applicant

APPENDIX A

CLAIMS ON APPEAL

1. (Previously Presented) For use in a system (100) capable of creating visual summaries of video material, an apparatus (130, 200) for creating a compact visual summary of video material, said apparatus (130, 200) comprising:

a visual summary controller (130, 200) capable of receiving keyframes of said video material;

wherein said visual summary controller (130, 200) is capable of extracting frame signatures from said keyframes, and capable of using said frame signatures to create superhistograms from said keyframes, and capable of using said frame signatures and said superhistograms to select representative keyframe images for each superhistogram to create a compact visual summary of said video material,

wherein said representative images include at least one of (1) the first image in each family histogram, (2) the most meaningful image in each superhistogram, (3) a randomly chosen image, and (4) an image that is closest to the cluster center.

2. (Original) The apparatus (130, 200) as claimed in Claim 1 wherein said visual summary controller (130, 200) is capable of filtering said keyframes and extracting frame signatures from said filtered keyframes before using said frame signatures to create said superhistograms to create a compact visual summary of said video material.

3. (Original) The apparatus (130, 200) as claimed in Claim 2 wherein said visual summary controller (130, 200) is capable of creating said compact visual summary of said video material by using said superhistograms to cluster said filtered keyframes, and by adding a representative keyframe from said clustered keyframes to said compact visual summary of said video material.

4. (Original) The apparatus (130, 200) as claimed in Claim 2 wherein said frame signature is a histogram.

5. (Original) The apparatus (130, 200) as claimed in Claim 3 wherein the distance measure for clustering is equal to a histogram difference calculated by one of: L1 distance measure method, L2 distance measure method, histogram intersection method, Chi Square test method, and bin-wise histogram intersection method.

6. (Cancelled)

7. (Original) The apparatus (130, 200) as claimed in Claim 5 wherein said visual summary controller (130, 200) is capable of selecting a family histogram to use to create said compact visual summary of said video material.

8. (Original) The apparatus (130, 200) as claimed in Claim 1 wherein said visual summary controller (130, 200) further comprises:

a visual summary retrieval module (180) capable of retrieving a compact visual summary stored in a memory unit (120) and causing said compact visual summary to be displayed in response to a user request.

9. (Original) The apparatus (130, 200) as claimed in Claim 3 wherein said visual summary controller (130, 200) is capable of using said compact visual summary to access at least one portion of said video material.

10. (Original) The apparatus (130, 200) as claimed in Claim 3 wherein said visual summary controller (130, 200) is capable of using said compact visual summary to create new video material.

11. (Previously Presented) A system (100) capable of creating visual summaries of video material, said system (100) comprising an apparatus (130, 200) for creating a compact visual summary of video material, said apparatus (130, 200) comprising:

a visual summary controller (130, 200) capable of receiving keyframes of said video material;

wherein said visual summary controller (130, 200) is capable of extracting frame signatures from said keyframes, and capable of using said frame signatures to create superhistograms from said keyframes, and capable of using said frame signatures and said superhistograms to select representative keyframe images for each superhistogram to create a compact visual summary of said video material ,

wherein said representative images include at least one of (1) the first image in each family histogram, (2) the most meaningful image in each superhistogram, (3) a randomly chosen image, and (4) an image that is closest to the cluster center.

12. (Original) The system (100) as claimed in Claim 11 wherein said visual summary controller (130, 200) is capable of filtering said keyframes and extracting frame signatures from said filtered keyframes before using said frame signatures to create said superhistograms to create a compact visual summary of said video material.

13. (Original) The system (100) as claimed in Claim 12 wherein said visual summary controller (130, 200) is capable of creating said compact visual summary of said video material by using said superhistograms to cluster said filtered keyframes, and by adding a representative keyframe from said clustered keyframes to said compact visual summary of said video material.

14. (Original) The system (100) as claimed in Claim 12 wherein said frame signature is a histogram.

15. (Original) The system (100) as claimed in Claim 13 wherein the distance measure for clustering is equal to a histogram difference calculated by one of: L1 distance measure method, L2 distance measure method, histogram intersection method, Chi Square test method, and bin-wise histogram intersection method.

16. (Cancelled)

17. (Original) The system (100) as claimed in Claim 16 wherein said visual summary controller (130, 200) is capable of selecting a family histogram to use to create said compact visual summary of said video material.

18. (Original) The system (100) as claimed in Claim 11 wherein said visual summary controller (130, 200) further comprises:

a visual summary retrieval module (180) capable of retrieving a compact visual summary stored in a memory unit (120) and causing said compact visual summary to be displayed in response to a user request.

19. (Original) The system (100) as claimed in Claim 13 wherein said visual summary controller (130, 200) is capable of using said compact visual summary to access at least one portion of said video material.

20. (Original) The system (100) as claimed in Claim 13 wherein said visual summary controller (130, 200) is capable of using said compact visual summary to create new video material.

21. (Previously Presented) For use in a system (100) capable of creating visual summaries of video material, a method for creating a compact visual summary of video material, said method comprising the steps of:

receiving in a visual summary controller (130, 200) keyframes of said video material;

extracting frame signatures from said keyframes;

using said frame signatures to create superhistograms from said keyframes; and

using said frame signatures and said superhistograms to select representative keyframe images for each superhistogram to create a compact visual summary of said video material ,

wherein said representative images include at least one of (1) the first image in each family histogram, (2) the most meaningful image in each superhistogram, (3) a randomly chosen image, and (4) an image that is closest to the cluster center.

22. (Original) The method as claimed in Claim 21 further comprising the steps of:

filtering said keyframes received in said visual summary controller (130, 200); and extracting frame signatures from said filtered keyframes before using said frame signatures to create said superhistograms to create a compact visual summary of said video material.

23. (Original) The method as claimed in Claim 22 further comprising the steps of:

using said histograms to cluster said filtered keyframes; and

adding a representative keyframe from said clustered keyframes to said compact visual summary of said video material.

24. (Original) The method as claimed in Claim 23 wherein the distance measure for clustering is equal to a histogram difference calculated by one of: L1 distance measure method, L2 distance measure method, histogram intersection method, Chi Square test method, and bin-wise histogram intersection method.

25. (Cancelled)

26. (Original) The method as claimed in Claim 23 further comprising the step of:

selecting a family histogram to use to create said compact visual summary of said video material.

27. (Original) The method as claimed in Claim 23 further comprising the steps of:

retrieving a compact visual summary stored in a memory unit (120); and

causing said compact visual summary to be displayed in response to a user request.

28. (Original) The method as claimed in Claim 23 further comprising the step of:

causing said visual summary controller (130, 200) to use said compact visual summary to access at least one portion of said video material.

29. (Original) The method as claimed in Claim 23 further comprising the step of:

causing said visual summary controller (130, 200) to use said compact visual summary to create new video material.

30. (Previously Presented) For use in a system (100) capable of creating visual summaries of video material, computer-executable instructions stored on a computer-readable storage medium (125) for creating a compact visual summary of video material, the computer-executable instructions comprising the steps of:

receiving in a visual summary controller (130, 200) keyframes of said video material;

extracting frame signatures from said keyframes;

using said frame signatures to create superhistograms from said keyframes; and

using said frame signatures and said superhistograms to select representative keyframe images for each superhistogram to create a compact visual summary of said video material,

wherein said representative images include at least one of (1) the first image in each family histogram, (2) the most meaningful image in each superhistogram, (3) a randomly chosen image, and (4) an image that is closest to the cluster center.

31. (Original) The computer-executable instructions stored on a computer-readable storage medium (125) as claimed in Claim 30 further comprising the step of:

filtering said keyframes received in said visual summary controller (130, 200); and extracting frame signatures from said filtered keyframes before using said frame signatures to create said superhistograms to create a compact visual summary of said video material.

32. (Original) The computer-executable instructions stored on a computer-readable storage medium (125) as claimed in Claim 31 further comprising the steps of:

using said histograms to cluster said filtered keyframes; and

adding a representative keyframe from said clustered keyframes to said compact visual summary of said video material.

33. (Original) The computer-executable instructions stored on a computer-readable storage medium (125) as claimed in Claim 32 wherein the distance measure for clustering is equal to a histogram difference calculated by one of: L1 distance measure method, L2 distance measure method, histogram intersection method, Chi Square test method, and bin-wise histogram intersection method.

34. (Cancelled)

35. (Original) The computer-executable instructions stored on a computer-readable storage medium (125) as claimed in Claim 34 further comprising the step of:

selecting a family histogram to use to create said compact visual summary of said video material.

36. (Original) The computer-executable instructions stored on a computer-readable storage medium (125) as claimed in Claim 30 further comprising the steps of:

retrieving a compact visual summary stored in a memory unit (120);

and causing said compact visual summary to be displayed in response to a user request.

37. (Original) The computer-executable instructions stored on a computer-readable storage medium (125) as claimed in Claim 32 further comprising the step of:

causing said visual summary controller (130, 200) to use said compact visual summary to access at least one portion of said video material.

38. (Original) The computer-executable instructions stored on a computer-readable storage medium (125) as claimed in Claim 32 further comprising the step of:

causing said visual summary controller (130, 200) to use said compact visual summary to create new video material.

APPENDIX B

Evidence on Appeal

None

APPENDIX C

Related Proceedings on Appeal

None